Research Techniques Made Simple: Using Genetic Variants for Randomization

Short title: Assessing causality with Mendelian Randomization

Ashley Budu-Aggrey[1] and Lavinia Paternoster[1]

[1]Medical Research Council (MRC) Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

Corresponding author: Lavinia Paternoster

Address: Medical Research Council (MRC) Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

Telephone: +44 (0)117 33 10135

Email: L.Paternoster@bristol.ac.uk

Author roles

Ashley Budu-Aggrey: Research Associate (Email: ashley.budu-aggrey@bristol.ac.uk)

Lavinia Paternoster: Senior Lecturer in Genetic Epidemiology (Email: L.Paternoster@bristol.ac.uk)

Keywords

Epidemiology, Genetics, Inflammatory skin diseases, Statistics

Abbreviations

25OHD	25-hydroxyvitamin D
2SLS	Two-Stage Least Squares
AD	Atopic Dermatitis
BMI	Body Mass Index
eQTLs	Expression Quantitative Trait Loci

GRS	Genetic risk score
GWAS	Genome-Wide Association Study
HUNT	Nord-Trøndelag Health Study
IV	Instrumental Variable
IVW	Inverse-Variance Weighted
mQTL	Methylation Quantitative Trait Loci
MR	Mendelian Randomization
pQTL	Protein Quantitative Trait Loci
RCT	Randomised Controlled Trial
SNP	Single-Nucleotide Polymorphism

This work was completed in Bristol, United Kingdom

Abstract

Observational epidemiological studies have identified associations between a number of modifiable exposures and outcomes, including in dermatology, such as between smoking and psoriasis. However, it is challenging to determine if such relationships are causal, due to the potential of confounding and reverse causation. Mendelian Randomization (MR) is a statistical method which can be used to investigate the causal relationships between an exposure and outcome, by using a genetic instrument that proxies the exposure. The resulting estimate (under certain assumptions) can be interpreted as the causal estimate, free of confounding and reverse causation. In this review, we provide an overview of how to undertake an MR analysis, with examples from the dermatology literature. We also discuss the challenges and future directions of this method.

Glossary

Term	Description
Confounder	A variable that is a common cause of both the exposure and the outcome
Exposure	An explanatory variable used to explain or predict an outcome variable, such as a trait or disease
F statistic	Obtained from the regression of a response variable upon a predictor variable, for example the regression of the exposure of interest upon an instrumental variable (IV). This can be used as a measure of the strength of association between an IV and the exposure, hence giving an indication of the strength of the instrument. The further away the F statistic is from 1, the stronger the instrument. The F statistic also depends on the size of the sample.
GWAS	Genome-wide association study. Involves analysing genetic variants across the genome, such as single-nucleotide polymorphisms (SNPs) for association with a disease or trait of interest.

Instrumental variable (IV)	A variable that is associated with an exposure of interest, but not the outcome. In MR studies genetic variants are used as IVs. A valid IV must also be independent of confounders of the exposure-outcome association and must only affect the outcome via the exposure.
Mendelian Randomization	A method for assessing the causal effect of an exposure upon an outcome, using genetic variants as instruments or "proxies" for the exposure variable.
MR-base	A centralized database of summary GWAS data, and an analytical platform to perform Mendelian Randomization and sensitivity analyses
PheWAS	Phenome-wide association study. Involves analysing the association between genetic variants and multiple phenotypic variables (on a phenome-wide scale) rather than a single phenotype.
Pleiotropy	Occurs when a genetic instrument is independently associated with multiple risk factors for the outcome, in addition to the exposure of interest. This results in the third IV assumption being violated that assumes that the genetic instrument only affects the outcome via the exposure.
Reverse causality	Where an association is due to the assumed outcome variable influencing the exposure variable, rather than the exposure influencing the outcome
Sensitivity Analysis	Performed to assess the robustness of the main analysis, or the validity of the main results

Introduction

Observational epidemiological studies have uncovered relationships between disease and various explanatory factors known as exposures. Notable examples in dermatology include the association of psoriasis with smoking (Armstrong et al. 2014), and more recently the association of atopic dermatitis (AD) with cardiovascular traits (Standl et al. 2017). However, traditional observational studies are prone to biases such as confounding, where the observed association may be due to the exposure being related to other lifestyle or socioeconomic factors that have a casual influence on disease. Furthermore, observed associations may be due to reverse causation, where disease is actually influencing the assumed exposure (Lawlor et al. 2008), for example having psoriasis could influence an individual's propensity to smoke. Mendelian Randomization (MR) presents as a method to evaluate causality in an observational study setting. We aim to provide an overview of the principle of MR and the statistical methods used.

The principle of MR

MR is a form of instrumental variable (IV) analysis where genetic variants are used as instruments (or proxies) for an exposure of interest. As genetic variants are randomly segregated at conception and cannot be influenced by confounding factors or the outcome itself, they can be used to estimate the causal effect of the exposure upon an outcome (Lawlor et al. 2008) (Figure 1).

Performing MR requires two pieces of information. (1) the effect of the genetic instrument on the exposure (β_{XZ}), and (2) the effect of the genetic instrument on the outcome (β_{YZ}). These can then be used to estimate the causal effect of the exposure on the outcome ($Causal\ \beta_{YX}$) using the following ratio (Wald 1940):

$$Causal\ \beta_{YX} = \frac{\beta_{YZ}}{\beta_{XZ}}$$

Eq 1 - The Wald ratio can be used to estimate the casual effect (β_{YX}) of an exposure (*X*) upon an outcome (*Y*).

For a genetic variant to qualify as an IV, three core assumptions must be satisfied. The variants (1) must be truly associated with the exposure of interest; (2) must not be associated with confounders of the exposure-outcome relationship; and (3) must only affect the outcome via the exposure and not through an alternative pathway (Zheng et al. 2017). The use of genetic variants in an MR framework can be compared to an RCT, where genotypes are used to randomise individuals to different subgroups (Lawlor et al. 2008). The effect of the genetic instrument on the outcome (β_{YZ}) is analogous to an intention-to-treat effect from an association between randomisation and an outcome in an RCT (Burgess and Thompson 2015).

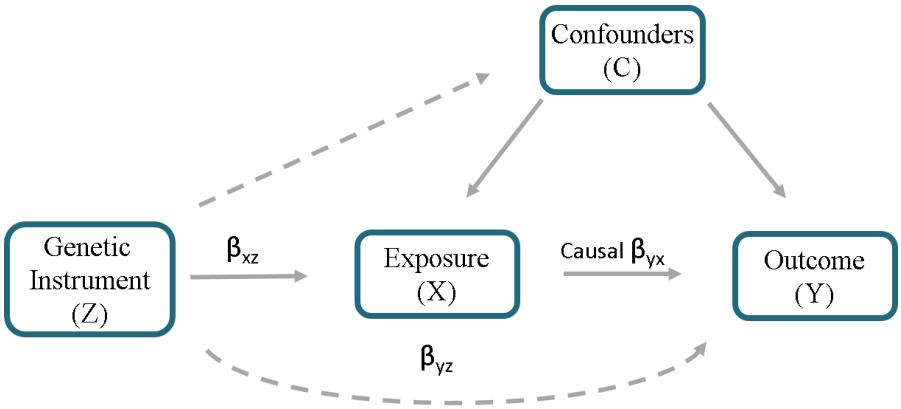

***Figure 1**- Illustrative diagram of standard Mendelian Randomization (MR) analysis. A valid genetic instrument (Z) must be truly associated with the exposure (X), must not be associated with confounders (C), and should only have an effect on the outcome (Y) via the exposure. Dashed arrows represent violations of these MR assumptions.*

As MR requires estimates of the associations between genetic variants and the exposure and genetic variants and the outcome, the rise of genome-wide association studies (GWAS) (Tsoi et al. 2018) provide a wealthy resource of genetic instruments for MR. Published summary GWAS data can be obtained from various sources such as the GWAS catalogue (www.ebi.ac.uk/gwas/), MR-base (www.mrbase.org), or direct from the authors of the GWAS (Figure 2). Commonly, independent SNPs that have been reported to be associated with an exposure on a genome-wide significant level (*P*-

value<$5x10^{-8}$) are used as genetic instruments for the exposure (Zheng et al. 2017), but MR analyses can be conducted using just a single genetic variant, or even using all variants in the genome (appropriately weighted by their effect on the exposure). Published MR studies in dermatology include those investigating causal relationships between fatty acids and melanoma (Liyanage et al. 2018), vitamin D levels and AD risk (Manousaki et al. 2017) as well as skin-aging (Noordam et al. 2017), and most recently with BMI and psoriasis risk (Budu-Aggrey et al. 2019) which will be referred to throughout this review.

MR approaches and statistical methods

MR study designs

A basic MR study design involves obtaining all information required from the same set of individuals, meaning the genetic, exposure and outcome data are all available from the same study. This is known as **one-sample MR**. Large population-based studies such as the UK Biobank provide ideal datasets for such analysis to be carried out. However, it may not always be possible to gather exposure and outcome measures from the same dataset. **Two-sample MR** is therefore more commonly adopted, where the effect of genetic variants on the exposure is obtained from one sample, and the effect of genetic variants on the outcome is obtained from another. This approach has been greatly facilitated by the increasing availability of summary GWAS data, as well as analytical platforms to perform two-sample MR, such as MR-base. The steps for a two-sample MR are shown in Figure 2 (Hemani et al. 2018).

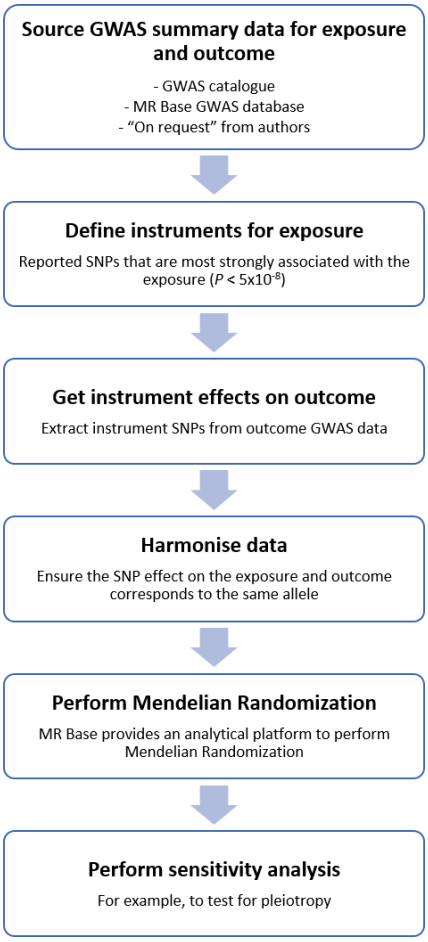

Figure 2- Workflow for performing two-sample MR analysis, adapted from Elife (Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-base platform supports systematic causal inference across the human phenome. Elife. 2018;7: e34408)

We recently investigated causality between BMI and psoriasis using both one-sample MR with individual-level data from the UK Biobank and Nord-Trøndelag Health Study (HUNT), and two-sample MR with published summary GWAS data. Consistent results were obtained from both analyses. The combined causal estimate suggested a 9% increase in the risk of psoriasis per 1 unit increase in BMI (Budu-Aggrey et al. 2019) (Figure 3). This finding supports previous reports of weight loss improving the prognosis of psoriasis (Maglio et al. 2017) and could suggest weight control as an intervention to prevent or treat psoriasis.

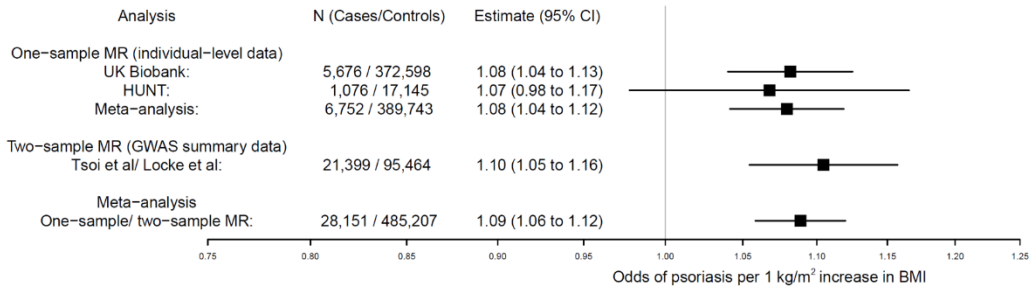

Figure 3 - One-sample and two-sample MR estimates give evidence of increased psoriasis risk with 1 unit increase in BMI (kg/m^2). Adapted from Public Library of Science: PLoS Medicine (Budu-Aggrey A, Brumpton B, Tyrrell J, Watkins S, Modalsli EH, Celis-Morales C, et al. Evidence of a causal relationship between body mass index and psoriasis: A mendelian randomization study. Lewis C, editor. PLOS Med. Public Library of Science; 2019;16(1):e1002739), © 2019.

A bi-directional MR approach can also be adopted, which investigates causal effects in both directions. This requires suitable genetic instruments to be available for both traits. Such analysis can help uncover the direction of causality that explains the observational association. For example, when considering the relationship between BMI and psoriasis, we performed bi-directional MR and found evidence that the observational relationship is largely due to the causal effect of higher BMI on psoriasis risk, rather than a causal effect of psoriasis influencing BMI (Budu-Aggrey et al. 2019).

MR statistical methods

The simplest method to perform MR involves dividing the effect of the genetic instrument on the outcome by the effect of the genetic instrument on the exposure. This is commonly termed as the "ratio of coefficients method" or the "Wald ratio method" (as shown in Eq. 1) and can be performed with either summarized or individual-level data (Burgess et al. 2017). Two-stage methods can also be applied, such as two-stage least squares (2SLS), as used in the BMI and psoriasis paper (Budu-Aggrey et al. 2019). This method involves regressing the exposure upon the genetic instruments, then regressing the outcome upon the genetically predicted values from the first regression, which allows

for the true standard error to be estimated. Additional MR methods have been previously discussed elsewhere (Burgess et al. 2015).

Combining multiple variants

Where multiple genetic instruments are available for an exposure, these can be combined into a genetic risk score (GRS) and used as a single instrument to perform MR (Zheng et al. 2017). Alternatively, an inverse-variance weighted (IVW) approach can be applied, whereby the ratio estimate from each independent genetic variant is combined using a fixed-effect meta-analysis model, where each variant is assumed to provide independent information, and the contribution of each variant is the inverse of the variance of its effect on the outcome (Burgess et al. 2013).

Sensitivity methods

One major potential problem with MR is when the genetic instrument affects the outcome through an alternative pathway that is distinct from the exposure of interest (termed pleiotropy), which violates the third assumption (as outlined above). Various sensitivity methods have been developed to detect and address pleiotropy including MR-Egger regression, weighted-median analysis the mode-based estimate and the latent causal variable method. These methods have different assumptions, but aim to estimate the true causal effect in the presence of modest levels of pleiotropy (O'connor and Price; Zheng et al. 2017).

Methods and approaches for MR analysis	**Description**
MR study design	

One-sample MR	Performed with genetic instruments, exposure and outcome data that have been measured in the same sample population.
Two-sample MR	The effect of the genetic instruments on exposure is obtained from a different sample to the effect of the genetic instruments on the outcome.
Bi-directional MR	The causal relationship between two traits is investigated in both directions. This approach can be applied to one-sample or two-sample MR.
Statistical methods	
Wald ratio method	Performed with a single genetic instrument (or genetic risk score) by dividing the coefficient of the outcome-instrument association by the coefficient of the exposure-instrument association.
Two-stage least squares (2SLS) regression	Involves two regression stages where the exposure is regressed upon the genetic instruments. The outcome is then regressed upon the genetically predicted exposure values from the first-stage regression.
Combining multiple variants	
Inverse-variance weighted (IVW) estimator	Combination of ratio estimates from individual variants in a fixed-effect meta-analysis. The

	contribution of each instrument is the inverse of the variance of its effect on the outcome.
Genetic risk score (GRS)	Multiple genetic instruments for an exposure are combined into a genetic risk score. This can then be used as a single instrument to perform MR.
Sensitivity analysis	
MR-Egger regression	Sensitivity analysis to perform MR with multiple instruments. This can be used to detect pleiotropy and provide a causal estimate that is robust to pleiotropy.
Weighted-median estimator	Sensitivity analysis to perform MR with multiple instruments. Will provide consistent causal estimates when at least 50% of the information in the analysis come from valid genetic instruments.
Mode-based estimator	An MR sensitivity analysis which will provide a robust causal estimate in the presence of pleiotropy, if the most common pleiotropy value is zero across the genetic instruments.
Latent causal variable analysis	Distinguishes between genetic correlation and causation by mediating the genetic correlation between two traits with a latent causal variable that itself has a causal effect on each trait.

Challenges and limitations of MR studies

Although MR has proven to be a useful tool to estimate causality, there are instances where MR may be limited, or the IV assumptions may be violated. In some cases, there may be only weak genetic instruments available for the exposure of interest. Genetic instruments that explain very little of the variance in exposure can result in weak instrument bias, where the causal estimates can be biased towards the null in a two-sample MR setting and towards the observational estimate in a one-sample MR setting (Zheng et al. 2017). This highlights the need for GWAS to uncover associated variants and strong, reliable instruments to perform MR. The F-statistic from the regression of the exposure on the genetic instrument indicates the strength of the instrument. It is recommended to use genetic variants with an F-statistic>10 (Burgess et al. 2013; Lawlor et al. 2008). As the F-statistic is dependent on sample size, weak instrument bias can also be addressed by utilising larger sample sizes (Burgess and Thompson 2015). Additionally, combining individual variants into a GRS increases the instrument strength. The instrument for BMI in our psoriasis analysis had an F-statistic of 7091, indicating a strong instrument for BMI (Budu-Aggrey et al. 2019).

Although it is assumed that a genetic instrument is independent of confounders, this cannot be tested for all potential confounders. However, it is sensible to test for association between the genetic instrument and any available measured potential confounders.

Applications and future directions for MR

MR is commonly performed to investigate the causality of established observational associations. However, a “hypothesis-free” approach can also be adopted to uncover novel causal relationships. This involves performing MR on a phenome-wide scale, known as MR-pheWAS where the effect of a single exposure upon multiple outcomes is evaluated. This has been demonstrated by Haycock et al,

where telomere length was found to increase the risk of several cancers, while reducing the risk of non-neoplastic diseases (Haycock et al. 2017).

MR can also be applied to investigate the causal role of molecular traits such as gene expression, methylation and protein biomarkers upon disease. In doing so, genetic variants associated with expression (eQTLs), methylation (mQTLs) or plasma protein levels (pQTLs) are used as genetic instruments for the exposure and can provide insight into the causal pathways that underlie disease. This has been demonstrated for AD, where MR analysis with pQTLS gave evidence that IL1RL2 and IL18R1 are causal proteins for AD risk (Sun et al. 2017).

Many MR studies are performed in cohorts with limited ethnic variation. As demonstrated by Ogawa et al, trans-ethnic MR studies can make the causal estimate more robust to confounding by population stratification and more generalisable to broader ethnic backgrounds (Ogawa et al. 2018).

We also expect that MR methods will begin to be applied to outcomes of disease progression (as opposed to onset), to enable them to be more informative for treatment of patients (Paternoster et al. 2017). Such studies have begun to emerge in other disease areas, such as Parkinson's disease (Simon et al. 2014), and could potentially uncover novel therapeutic targets or drug repurposing opportunities in dermatology.

Conclusion

MR has proven to be a robust statistical method to infer causal relationships in observational studies. In this review, we have presented strategies for performing MR, as well as the limitations and promising extensions of this method. As large GWAS summary statistics and open-access datasets become increasingly available, and additional methods continue to be developed, this will increase the potential for MR analysis to produce further evidence of causality for dermatological traits. This

in turn will aid the understanding of underlying mechanisms of disease and inform disease prevention and treatment.

Conflict of interest

LP has received personal fees from Merck for Scientific Input Engagement related to MR methodology.

Acknowledgements

AB-A is funded by a grant awarded by the British Skin Foundation (8010 Innovative Project), awarded to LP. AB-A, and LP work in a research unit funded by the UK Medical Research Council (MC_UU_00011/1).

Summary points

- Mendelian Randomization (MR) is a statistical method for investigating causality between exposure and outcome variables in observational epidemiology
- Unlike traditional observational studies, MR uses genetic variants as instruments (or proxies) for the exposure, hence avoiding confounding and reverse causation
- Application of such methods in the field of dermatology is a promising area of research
- Future directions and developments will allow MR to be valuable tool for investigating causal pathways for disease, as well as providing insight into therapeutic interventions

References

Armstrong AW, Harskamp CT, Dhillon JS, Armstrong EJ, Armstrong AW. Psoriasis and smoking: a

systematic review and meta-analysis Funding sources. Br. J. Dermatol. 2014;170:304–14 Available from: https://onlinelibrary.wiley.com/doi/pdf/10.1111/bjd.12670

Budu-Aggrey A, Brumpton B, Tyrrell J, Watkins S, Modalsli EH, Celis-Morales C, et al. Evidence of a causal relationship between body mass index and psoriasis: A mendelian randomization study. Lewis C, editor. PLOS Med. Public Library of Science; 2019;16(1):e1002739 Available from: http://dx.plos.org/10.1371/journal.pmed.1002739

Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet. Epidemiol. 2013;37(7):658–65 Available from: http://onlinelibrary.wiley.com/doi/10.1002/gepi.21758/abstract;jsessionid=92A157351FD56F4DE8FAE75C70F4730A.f03t03

Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. Stat. Methods Med. Res. 2015;0962280215597579 Available from: https://www.repository.cam.ac.uk/bitstream/handle/1810/248899/Burgess et al 2015 Statistical Methods of Medical Research.pdf?sequence=3

Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization [Internet]. Stat. Methods Med. Res. 2017 [cited 2018 Oct 1]. p. 2333–55 Available from: http://journals.sagepub.com/doi/pdf/10.1177/0962280215597579

Burgess S, Thompson SG. Mendelian Randomization: Methods for Using Genetic Variants in Causal Estimation [Internet]. ProtoView. CRC Press; 2015 [cited 2018 Oct 22]. Available from: https://books.google.co.uk/books?id=WYSbBgAAQBAJ&printsec=frontcover&dq=Mendelian+Randomization+Methods+for+using+Genetic+variants+in+Casual+Estimation+CRC+Press&hl=en&sa=X&ved=0ahUKEwjbwuT9k5reAhUqBMAKHYvmBlMQ6AEIMjAB#v=onepage&q&f=false

Haycock PC, Burgess S, Nounu A, Zheng J, Okoli GN, Bowden J, et al. Association Between Telomere Length and Risk of Cancer and Non-Neoplastic Diseases. JAMA Oncol. American Medical Association;

2017;3(5):636 Available from:
http://oncology.jamanetwork.com/article.aspx?doi=10.1001/jamaoncol.2016.5945

Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-base platform supports systematic causal inference across the human phenome. Elife. 2018;7 Available from: https://elifesciences.org/articles/34408

Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat. Med. 2008;27(8):1133–63 Available from: http://www.ncbi.nlm.nih.gov/pubmed/17886233

Liyanage UE, Law MH, Ong JS, Cust AE, Mann GJ, Ward S V., et al. Polyunsaturated fatty acids and risk of melanoma: A Mendelian randomisation analysis. Int. J. Cancer. John Wiley & Sons, Ltd; 2018;143(3):508–14 Available from: http://doi.wiley.com/10.1002/ijc.31334

Maglio C, Peltonen M, Rudin A, Carlsson LMS. Bariatric Surgery and the Incidence of Psoriasis and Psoriatic Arthritis in the Swedish Obese Subjects Study. Obesity (Silver Spring). Wiley-Blackwell; 2017;25(12):2068–73 Available from: http://www.ncbi.nlm.nih.gov/pubmed/29178583

Manousaki D, Paternoster L, Standl M, Moffatt MF, Farrall M, Bouzigon E, et al. Vitamin D levels and susceptibility to asthma, elevated immunoglobulin E levels, and atopic dermatitis: A Mendelian randomization study. Minelli C, editor. PLoS Med. Public Library of Science; 2017;14(5):e1002294 Available from: http://dx.plos.org/10.1371/journal.pmed.1002294

Noordam R, Hamer MA, Pardo LM, van der Nat T, Kiefte-de Jong JC, Kayser M, et al. No Causal Association between 25-Hydroxyvitamin D and Features of Skin Aging: Evidence from a Bidirectional Mendelian Randomization Study. J. Invest. Dermatol. Elsevier; 2017;137(11):2291–7 Available from: https://www.sciencedirect.com/science/article/pii/S0022202X17326829?via%3Dihub

O'connor LJ, Price AL. Distinguishing genetic correlation from causation across 52 diseases and complex traits. Available from: http://dx.doi.org/10.1101/205435

Ogawa K, Stuart PE, Tsoi LC, Suzuki K, Nair RP, Mochizuki H, et al. A trans-ethnic Mendelian randomization study identifies causality of obesity on risk of psoriasis. J. Invest. Dermatol. Elsevier; 2018; Available from: https://www.sciencedirect.com/science/article/pii/S0022202X18329154

Paternoster L, Tilling K, Davey Smith G. Genetic epidemiology and Mendelian randomization for informing disease therapeutics: Conceptual and methodological challenges [Internet]. Barsh GS, editor. PLoS Genet. Public Library of Science; 2017 [cited 2017 Oct 9]. p. e1006944 Available from: http://dx.plos.org/10.1371/journal.pgen.1006944

Simon KC, Eberly S, Gao X, Oakes D, Tanner CM, Shoulson I, et al. Mendelian randomization of serum urate and parkinson disease progression. Ann. Neurol. John Wiley & Sons, Ltd; 2014;76(6):862–8 Available from: http://doi.wiley.com/10.1002/ana.24281

Standl M, Tesch F, Baurecht H, Rodríguez E, Müller-Nurasyid M, Gieger C, et al. Association of Atopic Dermatitis with Cardiovascular Risk Factors and Diseases. J. Invest. Dermatol. Elsevier; 2017;137(5):1074–81 Available from: http://www.sciencedirect.com/science/article/pii/S0022202X16327919?via%253Dihub

Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al. Consequences Of Natural Perturbations In The Human Plasma Proteome. doi.org. Cold Spring Harbor Laboratory; 2017;134551 Available from: https://www.biorxiv.org/content/early/2017/05/05/134551

Tsoi LC, Patrick MT, Elder JT. Research Techniques Made Simple: Using Genome-Wide Association Studies to Understand Complex Cutaneous Disorders. J. Invest. Dermatol. Elsevier; 2018;138(3):e23–9 Available from: https://www.sciencedirect.com/science/article/pii/S0022202X18300216?via%3Dihub

Wald A. The Fitting of Straight Lines if Both Variables are Subject to Error. Ann. Math. Stat. Institute of Mathematical Statistics; 1940;11(3):284–300 Available from: http://projecteuclid.org/euclid.aoms/1177731868

Zheng J, Baird D, Borges M-C, Bowden J, Hemani G, Haycock P, et al. Recent Developments in Mendelian Randomization Studies. Curr. Epidemiol. Reports. Springer International Publishing; 2017;4(4):330–45 Available from: http://link.springer.com/10.1007/s40471-017-0128-6

Multiple choice questions

1. **Which of the following is a limitation of observational studies that can be addressed with MR?**
   A. Publication bias
   B. Selection bias
   C. Confounding
   D. Inadequate sample size

2. **Which of the following is NOT an assumption for a valid MR instrument?**
   A. The instrument must be truly associated with the exposure and the outcome
   B. The instrument must be truly associated with the exposure
   C. The instrument must not be associated with confounders of the exposure-outcome relationship
   D. The instrument must only affect the outcome via the exposure

3. **Which of the following can be used to uncover the direction of a causal relationship?**
   A. Two-sample MR
   B. Observational analysis
   C. One-sample MR
   D. Bi-directional MR

4. **Which of the following can be used to address pleiotropy in MR**
   A. Wald ratio method
   B. MR-Egger regression
   C. Inverse-variance weighted estimator
   D. Two-stage least squares

5. Which of the following statements is FALSE

A. MR can be performed in a hypothesis-free manner

B. MR estimates represent the effect of long-term exposures

C. Pleiotropic genetic instruments cannot be included in MR analyses

D. MR can be used to investigate the causal role of molecular phenotypes

Multiple choice questions

1. **Which of the following are limitations of observational studies that can be addressed with MR?**
    - Publication bias
    - Selection bias
    - Confounding

Explanation: Traditional observational studies are limited by confounding, reverse causation and measurement error. MR can be used to evaluate causality in observational studies while avoiding these limitations.

    - Inadequate sample size

2. **Which of the following is NOT an assumption for a valid MR instrument?**
    - The instrument must be truly associated with the exposure and the outcome

Explanation: A valid MR instrument must satisfy three core assumptions. The instrument must be truly associated with the exposure, must not be associated with confounders of the exposure-outcome relationship, and must only affect the outcome via the exposure and not through an alternative pathway.

    - The instrument must be truly associated with the exposure
    - The instrument must not be associated with confounders of the exposure-outcome relationship
    - The instrument must only affect the outcome via the exposure

3. **Which of the following can be used to uncover the direction of a causal relationship?**
    - Two-sample MR
    - Observational analysis
    - One-sample MR
    - Bi-directional MR

Explanation: Bi-directional MR involves investigating the causal effect of an exposure upon an outcome, as well evaluating the effect in the reverse direction of the outcome upon the exposure. In doing so, the direction of the causal relationship can be determined

4. **Which of the following can be used to address pleiotropy in MR**
    - Wald ratio method
    - MR-Egger regression

Explanation: MR-Egger regression can be performed to detect the presence of pleiotropy and also obtain a causal estimate that is robust to pleiotropy

    - Inverse-variance weighted estimator
    - Two-stage least squares

5. **Which of the following statements are FALSE**
    - MR can be performed in a hypothesis-free manner
    - MR estimates represent the effect of long-term exposures
    - Pleiotropic genetic instruments cannot be included in MR analyses

Explanation: MR can be performed on a phenome-wide scale to investigate the causal effect of a single exposure upon multiple outcomes with MR-pheWAS. MR estimates also represent the effect of long-term exposures rather than short-term interventions. In addition, MR can be extended to investigate the causal effect of molecular traits upon disease, where eQTLS, mQTLS or pQTLs are used as genetic instruments. Genetic instruments that are pleiotropic are not valid for MR analysis, however, MR methods have been developed to address pleiotropy, that allow for both unpleiotropic and pleiotropic variants to be included. These include MR-Egger regression, weighted-median analysis and the mode-based estimate.

- MR can be used to investigate the causal role of molecular phenotypes